Description

SYSTEM AND METHOD FOR ENCODING AND DECODING AN IMAGE USING BITSTREAM MAP AND RECORDING MEDIUM THEREOF

Technical Field

[1]   The present invention is related to a system and method for encoding and decoding images. More particularly, the present invention is related to a system and method for encoding and decoding a panoramic image using a bitstream map.

Background Art

[2]   During the past 10 years, multimedia technology has undergone a dramatic change: from analog to digital. A digital media compression technique has been applied to various user products including digital TVs, DVDs and MP3's, etc., and has reduced a distance between a user and a contents provider by employing international standards, MPEG-1, MPEG-2 and MPEG-4, to the digital media compression technique. In particular, the MPEG-4, started standardization from 1993, has introduced a graphic object compression conception and has become general media of a new concept.

[3]   Recently, applications using three-dimensional video are gaining high spotlight. Three-dimensional video is classified into three categories: a panoramic video, a multi-view video and a stereo video. The stereo video is purposed to present stereo effect by providing different images to the left and right eyes of a viewer. The multi-view video provides different views of an object or scene by capturing at different view directions. The panoramic video, which also requires multiple capturing of scene, provides wider angle of a scene by combining multiple images of adjacent scenes.

[4]   FIG. 1 illustrates a method of obtaining three-dimensional video. Referring to FIG. 1, images obtained from multiple cameras using parallel or divergent views are combined to generate panoramic image data, whereas images from the convergent view are used for multi-view image data.

[5]   FIG. 2 illustrates a general process of panoramic video handling. Referring to FIG. 2, the process includes an acquirement step, a processing step and a displaying step. The acquirement step acquires a plurality of images using a single camera or multiple cameras. The acquired multiple images are synthesized into a single image and this process is called a registration. Herein, a characteristic value of a camera is an important factor in the registration. The synthesized image is encoded, transmitted, and

stored during the processing step. A reproduced panoramic video is displayed on a planar or nonplanar display.

[6]   Important factors in the acquirement step for acquiring the panoramic video include the number of cameras used for obtaining images, three-dimensional geometry and orientation information on each camera. These factors are essential for obtaining high quality panoramic image.

[7]   FIG. 3 shows an example of a panoramic picture reproduced on a cylindrical display. The display can be designed as simple as a cylinder or a sphere, and it can be designed more complicated like as a polygonal mesh. A two-dimensional map is still used to represent pixels in a cylinder or a sphere. Accordingly, it is effective that a cylinder or sphere is constructed by combining multiple planes as patches shown in FIG. 3, in terms of the quantity of calculations.

[8]   Without regarding to the shapes of display, a panoramic picture requires a larger capacity compared to a single-view picture. Thus, efficient compression and video data streaming are important.

[9]   When acquisition and synthesis of images for generating a panoramic image are finished, the panoramic image is encoded so as to store or transmit The panoramic video compression is not much different from conventional video compression, except for information on a polygonal model on which the panoramic image is displayed However, a fundamental difference between the panoramic video and a general video is that the panoramic video requires a larger capacity and higher picture quality compared to the general video.

[10]   In the case of the panoramic video, a 'random access' functionality is an important factor together with high compression efficiency. The random access function enables to directly access a desired frame or region in an image sequence.

[11]   A general method of video compression includes an interframe coding method that encodes a current frame using data predicted from a previous frame in order to remove temporal redundancy between similar pictures in sequence. The interframe coding method increases compression efficiency but it is not advantageous in terms of random access.

[12]   A video system that requires random access should accept delay time corresponding to at least several frames in order to access an interframe-coded image. In particular, in the case of the panoramic video, the delay time represents a vast amount of cal-culations and large memory capacity because the panoramic video has high resolution. In other words, the panoramic video has a problem that compression efficiency and

system complexity collide with each other.

[13]   Although the interframe coding method is replaced with an intraframe coding method in order to avoid the delay time corresponding to several frames, the delay time still remains depending on circumstances, for example, a period of time required for decoding a single frame can be problematical. A period of time required for decoding a single frame of a panoramic video having the size of 4Kx4K is estimated to be 120 times a period of time required for decoding a frame in SIF (352x240) format. Accordingly, real-time decoding may be difficult to be implemented since this calculation may take too much time even if a plurality of decoders is simultaneously employed thereto.

[14]   FIG. 4 illustrates an example of reproduction of a local area through a partial decoding in random access. Unless a special method is designed for local region decoding, the whole panoramic image must be decoded first even it is displaying a part of the whole image, and it is wasting system resources.

[15]   For example, let's assume the case of encoding a panoramic image by the conventional compression techniques such as JPEG, JPEG1200, MPEG4 Video SP Intra coding, MPEG4 Video SP Inter coding, MPEG4 AVC Intra coding and MPEG4 AVC Inter coding. Because these techniques were designed without considering the functionality of local region decoding, when we display the panoramic image on a smaller display screen, the whole image must be decoded first, then stored in the memory of devices, and finally the region data is sent to display.

[16]   Therefore, no aforementioned codec satisfies both the compression efficiency and the random access functionality. Thus, a new compression method satisfying these two functions is required to effectively compress the panoramic video.

Disclosure of Invention

Technical Problem

[17]   It is an advantage of the present invention to provide a system and method for encoding and decoding a panoramic image, which can process random access at a high speed while minimizing a decrease in compression efficiency in compression of the panoramic image.

Technical Solution

[18]   In one aspect of the present invention, an image compression system comprises an image segmentation unit for segmenting a first image into a plurality of sub-images; a first encoding unit for encoding the sub-images to output sub-image bitstreams; a

BMAP construction unit for calculating the quantity of information of each sub-image bitstream and generating BMAP information using the calculated quantity of information and information on construction of each sub-image; and a bitstream combining unit for combining the sub-image bitstreams and BMAP information.

[19]   In another aspect of the present invention, an image decoding system comprises a bitstream receiving unit for receiving bitstreams of a first image including a plurality of sub-images; a BMAP reading unit for reading BMAP information included in the bitstreams and outputting information on a sub-image to be decoded among the plurality of sub-images; a sub-image extracting unit for extracting a bitstream corresponding to the sub-image to be decoded from the bitstreams using the information on the sub-image to be decoded; and a sub-image decoding unit for decoding the bitstream extracted by the sub-image extracting unit.

[20]   In another aspect of the present invention, an image compressing method comprises receiving a first image; segmenting the first image into a plurality of sub-images; encoding the sub-images to generate sub-image bitstreams; calculating the quantity of information of each sub-image and generating BMAP information using the calculated quantity of information and information on construction of each sub-image; combining the sub-image bitstreams and the BMAP information to generate frame bitstreams; and combining the frame bitstreams to form the bitstream of the first image.

[21]   In another aspect of the present invention, an image decoding method comprises receiving bitstreams of a first image including a plurality of sub-images; reading information on a sub-image corresponding to a region to be decoded among the plurality of sub-images using BMAP information included in the bitstreams; and extracting a bitstream corresponding to the sub-image to be decoded from the bitstreams.

[22]   In another aspect of the present invention, a recording medium storing an image compression program, comprises a function of receiving an image including at least one frame; a function of segmenting the image into a plurality of sub-images; a function of encoding the sub-images to generate sub-image bitstreams; a function of calculating the quantity of information of each sub-image bitstream and generating BMAP information using the calculated quantity of information and information on construction of each sub-image; and a function of combining the sub-image bitstreams and BMAP information to generate frame bitstreams.

[23]   In another aspect of the present invention, a recording medium storing an image decoding program, comprises a function of receiving bitstreams including a plurality of sub-images; reading information on a sub-image including a region to be decoded

among the plurality of sub-images using BMAP information included in the bitstreams; and a function of extracting a bitstream corresponding to the sub-image to be decoded from the bitstreams.

[24]

[25]

Brief Description of the Drawings

[26] The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate an embodiment of the invention, and, together with the description, serve to explain the principles of the invention:

[27] FIG. 1 shows a method of acquiring an image in a three-dimensional video processing;

[28] FIG. 2 shows panoramic video processing steps;

[29] FIG. 3 shows an example of a reproduced panoramic video;

[30] FIG. 4 shows an example of reproduction of a local region through partial decoding of random access;

[31] FIG. 5 is a block diagram of an image compression system according to an embodiment of the present invention;

[32] FIG. 6 shows a panoramic image, which is input to an image segmentation unit and segmented into 8 tile images by the image segmentation unit, according to an embodiment of the present invention;

[33] FIG. 7 is a block diagram of an image decoding system according to an embodiment of the present invention;

[34] FIG. 8 is a flow chart showing an image compression method using JPEG standard according to an embodiment of the present invention;

[35] FIG. 9 shows a process of binarizing BMAP information according to an embodiment of the present invention;

[36] FIG. 10 illustrates a bitstream of an image having a plurality of frames;

[37] FIG. 11 is a block diagram of an image compression system according to another embodiment of the present invention; and

[38] FIG. 12 is a block diagram of an image decoding system according to another embodiment of the present invention.

Best Mode for Carrying Out the Invention

[39] In the following detailed description, only the preferred embodiment of the invention has been shown and described, simply by way of illustration of the best

mode contemplated by the inventor(s) of carrying out the invention. As will be realized, the invention is capable of modification in various obvious respects, all without departing from the invention. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not restrictive. To clarify the present invention, parts which are not described in the specification are omitted, and parts for which similar descriptions are provided have the same reference numerals.

[40]   FIG. 5 is a block diagram of an image compression system according to an embodiment of the present invention.

[41]   Referring to FIG. 5, the image compression system includes an image segmentation unit 110, an encoding unit 120, a BMAP constructing unit 130, a binarization unit 140, and a bitstream combining unit 150. The image segmentation unit 110 segments an input panoramic image into a plurality of tile images using tile construction information. Here, the tile images represent sub images included in the whole image, and the tile construction information includes information on a segmentation form. In the case that each tile image has the same size, for instance, horizontal and vertical pixel values of each tile image become the tile construction information. FIG. 6 shows a panoramic image segmented into eight of tile images by the image segmentation unit 110.

[42]   The encoding unit 120 compresses each of the tile images segmented by the image segmentation unit 110 to encode the tile images. According to an embodiment of the present invention, the encoding unit 120 can compress the images using JPEG coding. In this case, the encoding unit 120 includes a DCT (discrete cosine transform) unit 121 for discrete-cosine-transforming the tile images, a quantization unit 122 for quantizing the output signal of the DCT unit 121, and an entropy encoder 123. The internal configuration and operation of the encoding unit 120 are well known in the art so that detailed explanation therefor is omitted

[43]   The BMAP construction unit 130 receives bitstreams of the tile images output from the encoding unit 120, calculates the quantity of information generated for each tile image and combines the quantity of generated information and the tile construction information to generate BMAP (bitstream map) information. BMAP information according to an embodiment of the present invention is as follows. The BMAP information includes header information, tile construction information and the quantity of bits generated for each tile image.

[44]   Table 1

Header in-formation	Tile con-struction in-formation	Quantity of bits generated for tile 1	.....	Quantity of bits generated for tile N

[45]   The header of the BMAP information can include a synchronization bit, a start signal, a version number and so on. The header can include a variety of information items, as necessary.

[46]   The binarization unit 140 binarizes the BMAP information outputted from the BMAP construction unit 130. The bitstream combining unit 150 combines the bitstreams of the tile images outputted from the encoding unit 120 and the bitstream of the BMAP information outputted from the binarization unit 140 to form a final bitstream as follows.

[47]   Table 2

Header in-formation	BMAP	Tile 1 encoded bitstream	.....	Tile N encoded bitstream

[48]   The header of the final bitstream can include various information items such as titles and color information on images. The bitstream of the BMAP information is placed prior to the tile image bitstream in the final bitstream according to the embodiment, but the position of the BMAP information can be changed.

[49]   A decoding system receives the bitstream and reads the BMAP information from the received bitstream such that the position of a tile image including a desired region can be easily recognized.

[50]   Accordingly, an image only in the desired region can be decoded and provided to the user. Consequently, the present invention can rapidly provide a user-desired image with a small quantity of calculations whereas the conventional method that decodes the whole panoramic image, stores the decoded image and then provides an image in a desired region to the user.

[51]   A bitstream obtained by encoding a tile image is called "a tile image bitstream," and a bitstream obtained by combining the tile image bitstream with the BMAP information and then actually transmitted is called "a whole image bitstream," hereinafter.

[52]   FIG. 7 is a block diagram of an image decoding system according to an embodiment of the present invention. Referring to FIG. 7, the image decoding system includes a bitstream receiving unit 210, an interface unit 220, a BMAP reading unit

230, a tile extracting unit 240, and a tile decoding unit 250.

[53]   The bitstream receiving unit 210 receives a whole image bitstream transmitted from the image compression system and transmits the whole image bitstream to the BMAP reading unit 230. The interface unit 220 includes a keypad for receiving key signals and an interface used for communication with an external device. A user can select an image in a region to be decoded through the interface unit 220.

[54]   The BMAP reading unit 230 reads BMAP information from an input bitstream and extracts information on a tile image corresponding to the region to be decoded. Here, the information on the tile image includes the position of a bitstream corresponding to a tile to be decoded in the inputted whole image bitstreams (i.e., the number of bytes per tile). If required, the information on the tile image can include information on the size, scan direction/order and position of each tile. Here, the information on the position represents the position of a display image. That is, the BMAP reading unit 230 reads the BMAP information to detect a tile image to which the image on a region to be decoded belongs or corresponds (this process can be omitted in the case that tile construction information is already included in the decoding system such that the interface unit 220 can figure out which tile image to be decoded). In addition, the BMAP reading unit 230 reads the position of the bitstream corresponding to the tile image in the inputted bitstream.

[55]   The tile extracting unit 240 extracts the bitstream corresponding to the tile image using the information on the tile image, which is outputted from the BMAP reading unit 230. The tile decoding unit 250 decodes the extracted tile bitstream and provides the user with an image corresponding to the tile bitstream.

[56]   Accordingly, the decoding system according to an embodiment of the present invention can extract a tile image bitstream in a desired region and decode only the extracted tile image bitstream using the BMAP information, to thereby rapidly provide a user-desired image with a small quantity of calculations.

[57]   Various methods are used for compressing an input tile image according to embodiments of the present invention, and a method of compressing an image using JPEG will be explained below.

[58]   FIG. 8 is a flow chart showing a method of compressing an image using JPEG according to an embodiment of the present invention. Referring to FIG. 8, when an image is inputted in the step S301, the image segmentation unit 110 segments the image into a plurality of tiles in the step S302. Assume that the size of the input image is 1024x1024, and the image segmentation unit 110 segments the image into tiles each

of which has the size of 16x16. In other words, the image segmentation unit 110 segments the image into 16x16 tiles.

[59]   The encoding unit 120 JPEG-encodes the 16x16 tiles in the step S303. The JPEG-based compression technique can compress an image having an arbitrary size and carries out compression in a minimum unit of a 8x8 block. Thus, in the case of the 16x16 tile, there are four of 8x8 blocks.

[60]   The BMAP construction unit 130 calculates the quantity of information generated when the encoding unit 120 JPEG-encodes the tile images in the step S304. Here, the quantity of information generated when the tile images are encoded based on JPEG can be represented by a multiple of 8. Accordingly, the quantity of generated information can be measures in a unit of bytes and thus an actual numerical value can be a byte value generated for each tile. For instance, when the number of bits generated when an arbitrary tile is encoded is 256 bits, which become 32 bytes, and thus the quantity of generated information becomes 32 bytes.

[61]   The BMAP construction unit 130 constructs the BMAP information using the quantity of information generated for the tiles in the step S305. Here, since the size of each tile is defined as 16x16, information that must be included in the header information of each tile can include only the number of pixels in horizontal and vertical directions, that is, 16. In the case that the tiles are scanned in the horizontal direction first and then the vertical direction in sequence and this scanning order is defined in the image encoding system and decoding system, no additional information may be included in the tile construction information.

[62]   An example of BMAP information when the size of each tile is 16x16 is as follows.

[63]   Table 3

16	16	Number of bytes generated for Tile 1 (100 for example)	.....	Number of bytes generated for Tile N (150 for example)

[64]   In Table 3, the first two blocks represent header information.

[65]   The image segmentation unit 110 can segment an input image into tile images each of which has an arbitrary size. Thus, the BMAP information must include the width and the height of each tile image and information on a position of the respective corresponding tile image.

[66]   A method of finding a bitstream corresponding to a tile to be decoded in the whole

input bitstream using the BMAP information in the case that the BMAP information includes the number of bytes generated for each tile will now be explained

[67]   Assume that the quantity of information of the first tile is 100 bytes, the quantity of information of the second tile is 90 bytes, the quantity of information of the third tile is 130 bytes, the quantity of information of the fourth tile is 80 bytes, and the quantity of information of the Nth tile is 150 bytes. When only the fourth through tenth tiles are displayed, the bitstream corresponding to the third tile is finished at the 320th byte in the entire bitstream (320=100+90+130) and the bitstream corresponding to the fourth tile is finished at the 321th byte through the BMAP information. The finishing point of the tenth tile can be easily drawn, and further, random access can be carried out in a like manner.

[68]   The BMAP information generated by the BMAP construction unit 130 is represented by integers. Thus, the binarization unit 140 binarizes the BMAP information so as to be stored and transmitted in the step S306. A typical method of binarizing integers defines a range of maximum value and allocates a specific number of bits to the respective information. When 10 bits are allocated for the number of bytes generated for each tile image, for example, integers up to 1024 can be represented by binary values. However, this method may allocate the number of bits more than actually required. Furthermore, when an integer value larger than the maximum value is generated, additional information for processing the integer value must be transmitted. To solve this problem, unary arithmetic coding can be used. The unary arithmetic coding will now be explained

[69]   When 3, 4 and 2 are coded, for instance, 0s, as many as each value, are arranged, and 1 is finally arranged to binarize random natural number (i.e., 2 -> 001, 3 -> 0001, 4 -> 00001, 10 -> 0000000001).

[70]   According to another embodiment of the present invention, the binarization unit 140 can use differential pulse code modulation that obtains a difference between adjacent values in the BMAP information and binarizes the difference. When the number of bytes generated for three tiles are 10, 12 and 11, for example, 10, 2, and -1 are actually binarized and transmitted When applying the differential pulse code modulation to transmit the difference, a negative difference may be generated. Therefore, unary arithmetic coding is performed on the absolute value of the difference, and 1-bit sign information is added to before or after the absolute value. For example, when the unary arithmetic coded value is a positive number, 1 is added as sign information before the unary arithmetic coded value, and 0 is added before the

unary arithmetic coded value when the unary arithmetic coded value is a negative number. In this case, 2 is represented as 1001, and -2 is represented as 0001.

[71]   The binarization unit 140 binary-arithmetic-codes the binary numbers generated through the binarization of the BMAP information.

[72]   FIG. 9 shows a process of binarizing the BMAP information according to an embodiment of the present invention. As shown in FIG. 9, the header information and the number of bytes generated for each tile image are represented in binary forms and binary-arithmetic-coded to compress the BMAP information.

[73]   Here, the binary arithmetic coding process is purposed to reduce the quantity of information on data to be transmitted and, as necessary, the header information and the number of bytes generated for each tile image represented in binary forms can be transmitted without being subjected to the binary arithmetic coding process.

[74]   When the binarization of the BMAP information is accomplished, the bitstreams of the BMAP information and the bitstreams of the tile images are combined to form the whole image bitstream in the step S307.

[75]   Through the aforementioned process, a single panoramic image can be encoded using the BMAP information, and video information can be encoded when the process is repeated for multiple frames.

[76]   FIG. 10 illustrates a bitstream of an image having a plurality of frames.

[77]   As shown in FIG. 10, in the case of the image having multiple frames, it is possible to include BMAP information in the respective frame based on JPEG compression and transmit the image.

[78]   Hereinafter, a case of applying an image compression method to a spatial scalable video coding method will be described according to another embodiment of the present invention.

[79]   The spatial scalable video coding method is a type of motion picture compression method The spatial scalable video coding method, which transmits video information in an environment where channel capacity is varied, such as the Internet, is for the purpose of transmitting the maximum quantity of video information while adapting to a variation in network channel capacity. A typical standard of the spatial scalable video coding method is MPEG-4 Visual FGS (fine granular scalability).

[80]   FIG. 11 is a block diagram of an image compression terminal system according to another embodiment of the present invention.

[81]   Referring to FIG. 11, the image compression terminal system includes a down-sampling unit 410, a first encoding unit 420, an up-sampling unit 430, a summing unit

440 and a second encoding unit 450.

[82]   The down-sampling unit 410 down-samples an input image. Down-sampling reduces the width and the height of an image, and generally decreases the size of the image by 1/2 or 1/4.

[83]   The first encoding unit 420 interframe-codes the down-sampled image. In the case of MPEG-4 Visual FGS standard, the first encoding unit 420 encodes the image according to the conventional MPEG-4 Part2 Visual standard. A bitstream generated through this process is called "a base layer bitstream." The interframe coding method is well known in the art so that detailed explanation therefor is omitted.

[84]   The up-sampling unit 430 up-samples the base layer bitstream outputted from the first encoding unit 420. Up-sampling is a process of increasing the size of an image.

[85]   The summing unit 440 obtains a difference between the input image and the image outputted from the up-sampling unit 430 and outputs the difference as an enhancement layer image.

[86]   The second encoding unit 450 encodes the enhancement layer image outputted from the summing unit 440 to generate an enhancement layer image bitstream. The second encoding unit 450 can use a frame-wise compression method without using redundancy between the enhancement layer frames. In other words, compression of the second encoding unit 450 can be considered as an intra coding since the second encoding unit 450 encodes the image only by frames.

[87]   The second encoding unit 450 encodes the enhancement layer image such that the BMAP information can be included in the enhancement image. That is, the second encoding unit 450 segments the enhancement image into a plurality of tile images and combines the BMAP information, containing each tile construction information and the quantity of information generated for each tile image, and bitstreams of the tile images so as to output the enhancement layer bitstream.

[88]   Accordingly, a user terminal decoding bitstreams to provide an image to a user can read the BMAP information related to the enhancement layer image to decode only a minimum number of tile images to provide the whole image to the user. Therefore, the quantity of calculation can be effectively reduced while carrying out a process of decoding a high-resolution motion picture.

[89]   The down-sampled image is encoded through a process of the interframe coding according to the foregoing embodiment, other video compression methods can also be applied thereto as necessary.

[90]   FIG. 12 is a block diagram of a receiving terminal that receives an image bitstream

to provide an image according to another embodiment of the present invention.

[91]   Referring to FIG. 12, the receiving terminal includes a first processing unit 510 for receiving and decoding a base layer bitstream, a second processing unit 520 for receiving and decoding an enhancement layer bitstream, and a summing unit 530.

[92]   The first processing unit 510 includes a bitstream receiver 511 for receiving the base layer bitstream, a decoder 512 for decoding the received base layer bitstream, and an up-sampler 513 for up-sampling the decoded data.

[93]   The second processing unit 520 includes a bitstream receiver 521 for receiving the enhancement layer bitstream, an interface 522 providing a user interface, a BMAP reading unit 523 for reading BMAP information from the enhancement layer bitstream to extract information on a tile image required for decoding a desired region, a tile extraction unit 524 for receiving the information on the tile image to extract a tile image bitstream from the enhancement layer bitstream, and a decoder 525 for decoding the tile image bitstream.

[94]   The summing unit 530 sums up video signals outputted from the first and second processing units 510 and 520.

[95]   The receiving terminal having the above-described configuration can decode the whole base layer bitstream which is relatively small in size and, for the enhancement layer bitstream required to be decoded for providing high-resolution image, analyze the BMAP information to decode only a minimum number of tile images, to thereby provide the image of the desired region.

[96]   The present invention has been described in connection with what is presently considered to be the most practical and preferred embodiment, it is to be understood that the invention is not limited to the disclosed embodiments, but, on the contrary, is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

[97]   The processes according to the present invention can be embodied as computer readable codes on a computer readable recording medium. The computer readable recording medium can be any data storage device that can store data which can be thereafter read by a computer system. Examples of the computer readable recording medium include read-only memory (ROM), random-access memory (RAM), CD-ROMs, magnetic tapes, floppy disks, optical data storage devices, and carrier waves (i.e., transmission through the Internet). The computer readable recording medium can also be distributed over a network coupled computer system so that the computer readable code is stored and executed in a distributed fashion.

[98]

[99]